



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number:

10/667,966

Source:

IFWC

Date Processed by STIC:

7-2-04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER  
VERSION 4.2 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND  
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efb/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):  
U.S. Patent and Trademark Office, 220 20<sup>th</sup> Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby,  
Room 1B03, Arlington, VA 22202

Revised 05/17/04

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER:

10/667,966

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos     The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length     The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering     The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII     The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length     Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"     A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)     Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence  
 (2) INFORMATION FOR SEQ ID NO: X: (insert SEQ ID NO where "X" is shown)  
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: X: (insert SEQ ID NO where "X" is shown)  
 This sequence is intentionally skipped  
  
 Please also adjust the "(ii) NUMBER OF SEQUENCES" response to include the skipped sequences
- 8      Skipped Sequences  
    (NEW RULES)     Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence  
 <210> sequence id number  
 <400> sequence id number  
 000
- 9      Use of n's or Xaa's  
    (NEW RULES)     Use of n's and/or Xaa's have been detected in the Sequence Listing.  
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response     Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 X Use of <220>     Sequence(s) 1-15 missing the <220> "Feature" and associated numeric identifiers and responses.  
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
 (See "Federal Register," 07/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"     Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n/Xaa     "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

## RAW SEQUENCE LISTING

DATE: 07/02/2004

PATENT APPLICATION: US/10/667,966

TIME: 11:51:28

Input-Set : A:\Untitled.ST25.txt

Output Set: N:\CRF4\07022004\J667966.raw

3 <110> APPLICANT: Xie, Dong  
 4 Jiang, He  
 6 <120> TITLE OF INVENTION: Peptide Derivative Fusion Inhibitors of HIV Infection  
 8 <130> FILE REFERENCE: 63024.000002  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/667,966  
 C--> 10 <141> CURRENT FILING DATE: 2003-09-23  
 10 <150> PRIOR APPLICATION NUMBER: 60/412,797  
 11 <151> PRIOR FILING DATE: 2002-09-24  
 13 <160> NUMBER OF SEQ ID NOS: 15  
 15 <170> SOFTWARE: PatentIn version 3.1  
 17 <210> SEQ ID NO: 1  
 18 <211> LENGTH: 44  
 19 <212> TYPE: PRT  
 20 <213> ORGANISM: Artificial Sequence  
 22 <220> FEATURE:  
 23 <223> OTHER INFORMATION: Peptide sequence  
 25 <400> SEQUENCE: 1  
 27 Ser Leu Glu Gln Ile Trp Asn Asn Met Thr Trp Glu Glu Trp Asp Arg  
 28 1 5 10 15  
 31 Glu Ile Asn Asn Tyr Thr Glu Leu Ile His Glu Leu Ile Glu Glu Ser  
 32 20 25 30  
 35 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu  
 36 35 40  
 39 <210> SEQ ID NO: 2  
 40 <211> LENGTH: 34  
 41 <212> TYPE: PRT  
 42 <213> ORGANISM: Artificial Sequence  
 44 <220> FEATURE:  
 45 <223> OTHER INFORMATION: Peptide sequence  
 47 <400> SEQUENCE: 2  
 49 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Lys Leu Ile His  
 50 1 5 10 15  
 53 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu  
 54 20 25 30  
 57 Leu Leu  
 61 <210> SEQ ID NO: 3  
 62 <211> LENGTH: 39  
 63 <212> TYPE: PRT  
 64 <213> ORGANISM: Artificial Sequence  
 66 <220> FEATURE:  
 67 <223> OTHER INFORMATION: Peptide sequence  
 69 <400> SEQUENCE: 3  
 71 Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Leu Glu Gln Ala Gln

Does Not Comply  
 Corrected Diskette Needed

see item 11 on error  
 summary sheet

## RAW SEQUENCE LISTING

DATE: 07/02/2004

PATENT APPLICATION: US/10/667,966

TIME: 11:51:28

Input Set : A:\Untitled.ST25.txt

Output Set: N:\CRF4\07022004\J667966.raw

```

72 1           5           10           15
75 Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp
76           20           25           30
79 Ala Ser Leu Trp Glu Trp Phe
80           35
83 <210> SEQ ID NO: 4
84 <211> LENGTH: 36
85 <212> TYPE: PRT
86 <213> ORGANISM: Artificial Sequence
88 <220> FEATURE:
89 <223> OTHER INFORMATION: Peptide sequence
91 <400> SEQUENCE: 4
93 Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln
94 1           5           10           15
97 Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu
98           20           25           30
101 Trp Asn Trp Phe
102           35
105 <210> SEQ ID NO: 5
106 <211> LENGTH: 34
107 <212> TYPE: PRT
108 <213> ORGANISM: Artificial Sequence
110 <220> FEATURE:
111 <223> OTHER INFORMATION: Peptide sequence same error
113 <400> SEQUENCE: 5
115 Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His
116 1           5           10           15
119 Ser Leu Ile Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu
120           20           25           30
123 Leu Leu
127 <210> SEQ ID NO: 6
128 <211> LENGTH: 34
129 <212> TYPE: PRT
130 <213> ORGANISM: Artificial Sequence
132 <220> FEATURE:
133 <223> OTHER INFORMATION: Peptide sequence
135 <400> SEQUENCE: 6
137 Trp Gln Glu Trp Glu Arg Lys Val Asp Phe Leu Glu Glu Asn Ile Thr
138 1           5           10           15
141 Ala Leu Leu Glu Glu Ala Gln Ile Gln Gln Glu Lys Asn Met Tyr Glu
142           20           25           30
145 Leu Gln
149 <210> SEQ ID NO: 7
150 <211> LENGTH: 34
151 <212> TYPE: PRT
152 <213> ORGANISM: Artificial Sequence
154 <220> FEATURE:
155 <223> OTHER INFORMATION: Peptide sequence
157 <400> SEQUENCE: 7

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/667,966

DATE: 07/02/2004

TIME: 11:51:28

Input Set : A:\Untitled.ST25.txt

Output Set: N:\CRF4\07022004\J667966.raw

159 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Lys Leu Ile His  
 160 1 5 10 15  
 163 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Glu Asn Glu Gln Glu  
 164 20 25 30  
 167 Leu Leu

171 &lt;210&gt; SEQ ID NO: 8

172 &lt;211&gt; LENGTH: 44

173 &lt;212&gt; TYPE: PRT

174 &lt;213&gt; ORGANISM: Artificial Sequence

176 &lt;220&gt; FEATURE:

177 &lt;223&gt; OTHER INFORMATION: Peptide sequence

179 &lt;220&gt; FEATURE:

180 &lt;221&gt; NAME/KEY: MISC\_FEATURE

181 &lt;222&gt; LOCATION: (23)..(23)

182 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 moie

183 ty.

186 &lt;400&gt; SEQUENCE: 8

188 Ser Leu Glu Gln Ile Trp Asn Asn Met Thr Trp Glu Glu Trp Asp Arg

189 1 5 10 15

W--&gt; 192 Glu Ile Asn Asn Tyr Thr Xaa Leu Ile His Glu Leu Ile Glu Glu Ser

193 20 25 30

196 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu

197 35 40

200 &lt;210&gt; SEQ ID NO: 9

201 &lt;211&gt; LENGTH: 45

202 &lt;212&gt; TYPE: PRT

203 &lt;213&gt; ORGANISM: Artificial Sequence

205 &lt;220&gt; FEATURE:

206 &lt;223&gt; OTHER INFORMATION: Peptide sequence same error

208 &lt;220&gt; FEATURE:

209 &lt;221&gt; NAME/KEY: MISC\_FEATURE

210 &lt;222&gt; LOCATION: (45)..(45)

211 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 moie

212 ty.

215 &lt;400&gt; SEQUENCE: 9

217 Ser Leu Glu Gln Ile Trp Asn Asn Met Thr Trp Glu Glu Trp Asp Arg

218 1 5 10 15

221 Glu Ile Asn Asn Tyr Thr Glu Leu Ile His Glu Leu Ile Glu Glu Ser

222 20 25 30

W--&gt; 225 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Xaa

226 35 40 45

229 &lt;210&gt; SEQ ID NO: 10

230 &lt;211&gt; LENGTH: 34

231 &lt;212&gt; TYPE: PRT

232 &lt;213&gt; ORGANISM: Artificial Sequence

234 &lt;220&gt; FEATURE:

235 &lt;223&gt; OTHER INFORMATION: Peptide sequence

237 &lt;220&gt; FEATURE:

238 &lt;221&gt; NAME/KEY: MISC\_FEATURE

## RAW SEQUENCE LISTING

DATE: 07/02/2004

PATENT APPLICATION: US/10/667,966

TIME: 11:51:28

Input Set : A:\Untitled.ST25.txt

Output Set: N:\CRF4\07022004\J667966.raw

239 &lt;222&gt; LOCATION: (13)..(13)

240 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 moie

241 ty.

244 &lt;400&gt; SEQUENCE: 10

W--&gt; 246 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Xaa Leu Ile His

247 1 5 10 15

250 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Trp Glu

251 20 25 30

254 Leu Leu

258 &lt;210&gt; SEQ ID NO: 11

259 &lt;211&gt; LENGTH: 35

260 &lt;212&gt; TYPE: PRT

261 &lt;213&gt; ORGANISM: Artificial Sequence

263 &lt;220&gt; FEATURE:

264 &lt;223&gt; OTHER INFORMATION: Peptide sequence

266 &lt;220&gt; FEATURE:

267 &lt;221&gt; NAME/KEY: MISC\_FEATURE

268 &lt;222&gt; LOCATION: (35)..(35)

269 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 moie

270 ty.

273 &lt;400&gt; SEQUENCE: 11

275 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Glu Leu Ile His

276 1 5 10 15

279 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu

280 20 25 30

W--&gt; 283 Leu Leu Xaa

284 35

287 &lt;210&gt; SEQ ID NO: 12

288 &lt;211&gt; LENGTH: 39

289 &lt;212&gt; TYPE: PRT

290 &lt;213&gt; ORGANISM: Artificial Sequence

292 &lt;220&gt; FEATURE:

293 &lt;223&gt; OTHER INFORMATION: Peptide sequence same error

295 &lt;220&gt; FEATURE:

296 &lt;221&gt; NAME/KEY: MISC\_FEATURE

297 &lt;222&gt; LOCATION: (13)..(13)

298 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide  
 moie

299 ty.

302 &lt;400&gt; SEQUENCE: 12

W--&gt; 304 Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Leu Xaa Gln Ala Gln

305 1 5 10 15

308 Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp

309 20 25 30

312 Ala Ser Leu Trp Glu Trp Phe

313 35

316 &lt;210&gt; SEQ ID NO: 13

317 &lt;211&gt; LENGTH: 40

318 &lt;212&gt; TYPE: PRT

319 &lt;213&gt; ORGANISM: Artificial Sequence

## RAW SEQUENCE LISTING

DATE: 07/02/2004

PATENT APPLICATION: US/10/667,966

TIME: 11:51:28

Input Set : A:\Untitled.ST25.txt

Output Set: N:\CRF4\07022004\J667966.raw

```

321 <220> FEATURE:
322 <223> OTHER INFORMATION: Peptide sequence
324 <220> FEATURE:
325 <221> NAME/KEY: MISC_FEATURE
326 <222> LOCATION: (40)..(40)
327 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide
moie
328      ty.
331 <400> SEQUENCE: 13
333 Trp Gln Glu Trp Glu Gln Lys Ile Thr Ala Leu Ile Glu Gln Ala Gln
334 1          5          10          15
337 Ile Gln Gln Glu Lys Asn Glu Tyr Glu Leu Gln Lys Leu Asp Lys Trp
338          20          25          30
W--> 341 Ala Ser Leu Trp Glu Trp Phe Xaa
342          35          40
345 <210> SEQ ID NO: 14
346 <211> LENGTH: 34
347 <212> TYPE: PRT
348 <213> ORGANISM: Artificial Sequence
350 <220> FEATURE:
351 <223> OTHER INFORMATION: Peptide sequence same error
353 <220> FEATURE:
354 <221> NAME/KEY: MISC_FEATURE
355 <222> LOCATION: (13)..(13)
356 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide
moie
357      ty.
360 <400> SEQUENCE: 14
W--> 362 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Xaa Leu Ile His
363 1          5          10          15
366 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Glu Asn Glu Gln Glu
367          20          25          30
370 Leu Leu
374 <210> SEQ ID NO: 15
375 <211> LENGTH: 35
376 <212> TYPE: PRT
377 <213> ORGANISM: Artificial Sequence
379 <220> FEATURE:
380 <223> OTHER INFORMATION: Peptide sequence
382 <220> FEATURE:
383 <221> NAME/KEY: MISC_FEATURE
384 <222> LOCATION: (35)..(35)
385 <223> OTHER INFORMATION: Xaa represents a Lysine residue derivatized with a maleimide
moie
386      ty.
389 <400> SEQUENCE: 15
391 Trp Glu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Lys Leu Ile His
392 1          5          10          15
395 Glu Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Glu Asn Glu Gln Glu
396          20          25          30
W--> 399 Leu Leu Xaa
400          35

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/667,966

DATE: 07/02/2004  
TIME: 11:51:29

Input Set : A:\Untitled.ST25.txt  
Output Set: N:\CRF4\07022004\J667966.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; Xaa Pos. 23  
Seq#:9; Xaa Pos. 45  
Seq#:10; Xaa Pos. 13  
Seq#:11; Xaa Pos. 35  
Seq#:12; Xaa Pos. 13  
Seq#:13; Xaa Pos. 40  
Seq#:14; Xaa Pos. 13  
Seq#:15; Xaa Pos. 35

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/667,966

DATE: 07/02/2004

TIME: 11:51:29

Input Set : A:\Untitled.ST25.txt

Output Set: N:\CRF4\07022004\J667966.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:192 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:16  
L:225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:32  
L:246 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0  
L:283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:32  
L:304 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0  
L:341 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:32  
L:362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0  
L:399 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:32